

# Power Allocation Schemes for Multicell Massive MIMO Systems

Qi Zhang, Shi Jin, *Member, IEEE*, Matthew McKay, *Senior Member, IEEE*,  
David Morales-Jimenez, *Member, IEEE*, and Hongbo Zhu

**Abstract**—This paper investigates the sum-rate gains brought by power allocation strategies in multicell massive multiple-input multiple-output systems, assuming time-division duplex transmission. For both uplink and downlink, we derive tractable expressions for the achievable rate with zero-forcing receivers and precoders respectively. To avoid high complexity joint optimization across the network, we propose a scheduling mechanism for power allocation, where in a single time slot, only cells that do not interfere with each other adjust their transmit powers. Based on this, corresponding transmit power allocation strategies are derived, aimed at maximizing the sum rate per-cell. These schemes are shown to bring considerable gains over equal power allocation for practical antenna configurations (e.g., up to a few hundred). However, with fixed number of users ( $N$ ), these gains diminish as  $M \rightarrow \infty$ , and equal power allocation becomes optimal. A different conclusion is drawn for the case where both  $M$  and  $N$  grow large together, in which case: (i) improved rates are achieved as  $M$  grows with fixed  $M/N$  ratio, and (ii) the relative gains over the equal power allocation diminish as  $M/N$  grows. Moreover, we also provide applicable values of  $M/N$  under an acceptable power allocation gain threshold, which can be used as to determine when the proposed power allocation schemes yield appreciable gains, and when they do not. From the network point of view, the proposed scheduling approach can achieve almost the same performance as the joint power allocation after one scheduling round, with much reduced complexity.

**Index Terms**—Multicell, massive MIMO, power allocation.

## I. INTRODUCTION

Multiple-input multiple-output (MIMO) systems are now regarded as a key technology for cellular communication systems due to various advantages they offer over single-antenna systems, including transmit diversity, high data rates and reliability [1–3]. By transmitting parallel data streams, it is

known that the ergodic capacity of MIMO channels increases linearly with the minimum number of transmit and receive antennas [4]. The transmitted information streams are separated at the receiver by means of appropriate signal processing techniques. The optimum receivers and precoders lead to a complexity burden on the system implementation. Therefore, linear receivers and precoders such as those based on the zero-forcing (ZF) criteria are often considered, as they offer significantly lower complexity with tolerable performance [5].

Recently, multiuser MIMO (MU-MIMO) systems, in which the antenna array of the base station (BS) simultaneously serves a multiplicity of autonomous users in the same time-frequency resource, have attracted substantial interest. Compared with conventional MIMO systems, MU-MIMO can offer a spatial multiplexing gain even if the users have only a single antenna [6–10].

Though MU-MIMO has been supported in 3GPP standards, it can not achieve the transmission rates demanded in 5G. Therefore, the concept of massive MIMO, which considers the use of hundreds of antenna elements to serve tens of users simultaneously, has come to the forefront of wireless communications research [11]. Massive MIMO can reap all the benefits of conventional MIMO at a greater scale, whilst providing substantial improvements in energy and spectral efficiencies with low complexity [12–23]. In particular, if the number of BS antennas is far greater than the number of single-antenna users, the simplest linear receivers and precoders become optimal [14]. It is also revealed in [13] that, when the number of BS antennas grows without bound, the achievable rate can be improved to a considerable level and the uncorrelated noise, fast fading and intracell interference all vanish. Moreover, [15] demonstrates that the transmitted power in massive MIMO systems can be proportionally scaled down with the number of BS antennas while maintaining the same rate performance.

In power-limited systems, the equal power allocation policy cannot take full advantage of the energy resource and it is far from optimal. In this paper, allowing for different transmit powers at each user, we derive three tractable expressions for the achievable uplink rate—a lower bound, an upper bound and an approximation which lies between these two bounds. To avoid the joint optimization across the network, which involves high complexity algorithms with (potential) convergence issues, we proposed a scheduling mechanism for power allocation, where in a single time slot, only cells that do not interfere with each other adjust their transmit powers. Based on this, we subsequently derive power allocation strategies

Manuscript received November 15, 2014; revised March 30, 2015. This work was partly supported by the China 973 project under Grant 2013CB329005, the China 863 Program under Grant 2014AA01A705 and the Jiangsu Graduate Research and Innovation Project under Grant KYLXJ0806. The work of S. Jin was supported by the National Natural Science Foundation of China under Grant 61222102 and the International Science & Technology Cooperation Program of China under Grant 2014DFT10300. The work of M. McKay and D. Morales-Jimenez was supported by the Hong Kong Research Grants Council under grant number 616713.

Q. Zhang and H. Zhu are with Jiangsu Key Laboratory of Wireless Communications, Nanjing University of Posts and Telecommunications, Nanjing, 210003, P. R. China (email: zhangqiqi\_1212@126.com; zhuhb@njupt.edu.cn).

S. Jin is with the National Mobile Communications Research Laboratory, Southeast University, Nanjing, 210096, P. R. China (email: jin-shi@seu.edu.cn). S. Jin is the corresponding author.

M. McKay and D. Morales-Jimenez are with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong (email: eemckay@ust.hk; eedmorales@ust.hk).

which optimize each of these expressions and compare their relative performance. In addition, we present a tractable lower bound for the achievable downlink rate which is shown to be very tight, and similarly, a corresponding power allocation policy is derived. These schemes are shown to yield noticeable rate improvements over equal power allocation under various conditions. This is true, for example, for BSs housing up to a few hundred antennas ( $M$ ). As  $M$  increases, the gain brought by optimized power allocation becomes less significant, with the optimal assignment approaching uniform allocation as  $M \rightarrow \infty$  for fixed numbers of users ( $N$ ). To further characterize these effects more precisely, we also consider the regime where  $M$  and  $N$  grow large together, i.e., with  $M/N$  fixed, and show that: (i) improved rates are achieved as  $M$  and  $N$  both grow, and (ii) the relative gains over the equal power allocation depend on the ratio  $M/N$ , with larger gains for smaller ratios. By specifying a minimum required gain over the equal power assignment, we then provide the range of  $M/N$  ratios which satisfy the proposed constraint, and these results can be used to provide engineering insight into when the gains brought by optimized power allocation are sufficient to warrant the additional complexity relative to equal-power allocation, and when they are not. For example, for the uplink, if the total transmit power per cell is 20dB, systems with  $M/N < 12$  (i.e., less than 12 BS antennas per user) can obtain more than 10% rate gain with optimized power allocation. These observations indicate that while equal power allocation is asymptotically optimal in the theoretical massive MIMO regime (i.e., when  $M/N \rightarrow \infty$ ), for a wide range of practical scenarios in which the BS antennas is moderately larger than the number of users (but not infinitely larger), optimized power allocation can bring appreciable gains.

This paper is organized as follows. Section II describes the multicell MIMO system model for both uplink and downlink, and defines the associated achievable rate with ZF receivers and ZF precoders, respectively. In Section III, we present the lower and upper bound for the achievable uplink rate, followed by a more accurate approximation which lies between these bounds. Based on this approximation, the corresponding power allocation scheme is put forth. In Section IV, we analyze the achievable downlink rate and, based on a tractable lower bound which we derive, an effective power allocation scheme is proposed. In Section V, we provide a set of numerical results, while Section VI summarizes the main results of this paper.

*Notation*—Throughout the paper, vectors are expressed in lowercase boldface letters while matrices are denoted by uppercase boldface letters. We use  $\mathbf{X}^H, \mathbf{X}^T, \mathbf{X}^*$  and  $\mathbf{X}^{-1}$  to denote the conjugate-transpose, transpose, conjugate and inverse of  $\mathbf{X}$ , respectively. Moreover,  $\mathbf{I}_N$  denotes an  $N \times N$  identity matrix, and  $[\mathbf{X}]_{ij}$  is the  $(i, j)$ th entry of  $\mathbf{X}$ . Finally,  $\mathbb{E}\{\cdot\}$  is the expectation operator,  $\|\cdot\|$  is the Euclidean norm,  $(a)^+$  denotes  $\max\{a, 0\}$  and  $\text{tr}(\cdot)$  is the trace operation.

## II. SYSTEM MODEL

Consider a multicell scenario with a central cell, denoted as cell 0, and  $L$  interference cells surrounding it, denoted as cell 1 to  $L$ . Each cell contains a MU-MIMO system with  $N$  single-

antenna users and one  $M$ -antenna BS. Channel reciprocity is exploited between uplink and downlink.

### A. Uplink

On the uplink, it is assumed that all users simultaneously transmit data streams to their BSs in the same time-frequency resource and the transmit power for each user may be different. Let  $\mathbf{G}_{il}$  ( $i, l = 0, 1, \dots, L$ ) denote the  $M \times N$  MIMO channel matrix between the  $N$  users in the  $l$ th cell and the  $M$  BS antennas in the  $i$ th BS. Therefore, the  $M \times 1$  received vector at the BS of the central cell can be written as

$$\mathbf{y}_0^{\text{ul}} = \mathbf{G}_{00}\mathbf{\Omega}_0^{1/2}\mathbf{x}_0 + \sum_{l=1}^L \mathbf{G}_{0l}\mathbf{\Omega}_l^{1/2}\mathbf{x}_l + \mathbf{n}_0^{\text{ul}}, \quad (1)$$

where the  $N \times N$  diagonal matrix  $\mathbf{\Omega}_l$  contains  $[p_{1l}^{\text{ul}}, \dots, p_{nl}^{\text{ul}}, \dots, p_{Nl}^{\text{ul}}]$  along its main diagonal, while  $p_{nl}^{\text{ul}}$  is the transmitted power of the  $n$ th user in the  $l$ th cell,  $\mathbf{x}_l \in \mathbb{C}^{N \times 1}$  denotes the transmitted signal vector from all users in the  $l$ th cell, and  $\mathbf{n}_0^{\text{ul}} \in \mathbb{C}^{M \times 1}$  represents the vector of additive white Gaussian noise with entries having zero mean and unit variance.

We denote the independent channel coefficient between the  $n$ th user in the  $l$ th cell and the  $m$ th antenna of the  $i$ th BS as  $g_{minl} = [\mathbf{G}_{il}]_{mn}$ , which accounts for independent fast fading, geometric attenuation and log-normal shadow fading and can be expressed as [13]

$$g_{minl} = h_{minl}\sqrt{\beta_{inl}}, \quad (2)$$

where  $h_{minl}$  is the fast fading element from the  $n$ th user in the  $l$ th cell to the  $m$ th antenna of the  $i$ th BS, which has independent real and imaginary parts with zero mean and variance  $1/2$ . The large scale fading  $\beta_{inl}$  from the  $n$ th user in the  $l$ th cell to the  $i$ th BS models both the geometric attenuation and shadow fading. It is reasonable to assume the large scale fading coefficient to be constant across the antenna array, since the distance between users and the BS is much larger than the distance between antennas, and the value of  $\beta_{inl}$  changes very slowly with time. Then,

$$\mathbf{G}_{0l} = \mathbf{H}_{0l}\mathbf{D}_{0l}^{1/2}, \quad (3)$$

where  $\mathbf{H}_{0l}$  denotes the  $M \times N$  fast fading matrix between the users in the  $l$ th cell and the central cell's BS, i.e.,  $[\mathbf{H}_{0l}]_{mn} = h_{m0nl}$  and  $\mathbf{D}_{0l}$  is a  $N \times N$  diagonal matrix with  $[\mathbf{D}_{0l}]_{nn} = \beta_{0nl}$ . We assume all BSs have perfect channel state information (CSI) of channels in their own cell, which is a reasonable approximation in environments with low or moderate mobility. Linear processing is assumed at all receivers. Let  $\mathbf{A}_{00}$  be the  $M \times N$  linear receiver matrix of the central cell's BS, which depends on the channel matrix  $\mathbf{G}_{00}$ . The received signal vector at the BS is processed as

$$\mathbf{r}_0 = \mathbf{A}_{00}^H \mathbf{y}_0^{\text{ul}}. \quad (4)$$

For ZF receivers,  $\mathbf{A}_{00} = \mathbf{G}_{00} (\mathbf{G}_{00}^H \mathbf{G}_{00})^{-1}$ . Substituting this along with (1) into (4) gives

$$\mathbf{r}_0 = \mathbf{\Omega}_0^{1/2}\mathbf{x}_0 + \sum_{l=1}^L \mathbf{A}_{00}^H \mathbf{G}_{0l} \mathbf{\Omega}_l^{1/2} \mathbf{x}_l + \mathbf{A}_{00}^H \mathbf{n}_0^{\text{ul}}. \quad (5)$$

The  $n$ th element of  $\mathbf{r}_0$  can be further expressed as

$$r_{n0} = \sqrt{p_{n0}^{u1}} x_{n0} + \sum_{l=1}^L \sum_{c=1}^N \sqrt{p_{cl}^{u1}} \mathbf{a}_{0n0}^H \mathbf{g}_{0cl} x_{cl} + \mathbf{a}_{0n0}^H \mathbf{n}_0^{u1}, \quad (6)$$

where  $\mathbf{a}_{0n0}$  is the  $n$ th column of  $\mathbf{A}_{00}$ ,  $\mathbf{g}_{0cl}$  is the  $c$ th column of  $\mathbf{G}_{0l}$ , and  $x_{cl}$  denotes the  $c$ th element of the signal vector  $\mathbf{x}_l$  ( $c = 1, \dots, N; l = 1, \dots, L$ ). The desired signal in (6) is  $\sqrt{p_{n0}^{u1}} x_{n0}$ , and the remaining terms are considered as interference and noise. The interference-plus-noise term is modeled as Gaussian noise, which constitutes a “worst case” noise assumption [24]. The ergodic achievable uplink rate of the  $n$ th user in the central cell is then given as follows [15]

$$R_{n0}^{u1} = \mathbb{E} \left\{ \log_2 \left( 1 + \frac{p_{n0}^{u1}}{\sum_{l=1}^L \sum_{c=1}^N p_{cl}^{u1} |\mathbf{a}_{0n0}^H \mathbf{g}_{0cl}|^2 + \|\mathbf{a}_{0n0}\|^2} \right) \right\}. \quad (7)$$

### B. Downlink

On the downlink, we assume that the transmit power sent by the BS for each user may be different. Then, the received signal by users in the central cell can be given as

$$\mathbf{y}_0^{d1} = \mathbf{G}_{00}^T \mathbf{B}_{00} \Psi_0^{1/2} \mathbf{q}_0 + \sum_{l=1}^L \mathbf{G}_{l0}^T \mathbf{B}_{ll} \Psi_l^{1/2} \mathbf{q}_l + \mathbf{n}_0^{d1}, \quad (8)$$

where  $\mathbf{B}_{ll}$  ( $l = 1, \dots, L$ ) is the  $M \times N$  precoding matrix applied at the  $l$ th BS, which is dependent on the channel matrix  $\mathbf{G}_{ll}$ ,  $\Psi_l$  is a  $N \times N$  diagonal power matrix containing  $[p_{1l}^{d1}, \dots, p_{nl}^{d1}, \dots, p_{Nl}^{d1}]$  along its main diagonal, while  $p_{nl}^{d1}$  is the transmitted power of the  $l$ th BS intended for the  $n$ th user. Moreover,  $\mathbf{q}_l \in \mathbb{C}^{N \times 1}$  is the vector of signals intended for the  $N$  users in the  $l$ th cell and  $\mathbf{n}_0^{d1}$  contains white complex Gaussian noise with entries having zero mean and unit variance. To satisfy the power constraint at the BS,  $\mathbf{B}_{ll}$  is chosen such that [25]

$$\mathbb{E} \{ \text{tr} (\mathbf{B}_{ll} \mathbf{B}_{ll}^H) \} = 1. \quad (9)$$

We assume users have perfect CSI and detect received signals with the optimal maximum likelihood receivers. The ZF linear precoder is given by

$$\mathbf{B}_{ll} = \alpha_l \mathbf{G}_{ll}^* (\mathbf{G}_{ll}^T \mathbf{G}_{ll}^*)^{-1}, \quad (10)$$

where the constant scalar  $\alpha_l$  is chosen to conform to the constraint (9). The substitution of (10) into (8) yields

$$\mathbf{y}_0^{d1} = \alpha_0 \Psi_0^{1/2} \mathbf{q}_0 + \sum_{l=1}^L \mathbf{G}_{l0}^T \mathbf{B}_{ll} \Psi_l^{1/2} \mathbf{q}_l + \mathbf{n}_0^{d1}, \quad (11)$$

and the  $n$ th element of  $\mathbf{y}_0^{d1}$  can be further expressed as

$$y_{n0}^{d1} = \alpha_0 \sqrt{p_{n0}^{d1}} q_{n0} + \sum_{l=1}^L \sum_{c=1}^N \sqrt{p_{cl}^{d1}} \mathbf{g}_{ln0}^T \mathbf{b}_{lcl} q_{lc} + n_{n0}^{d1}, \quad (12)$$

where  $\mathbf{g}_{ln0}$  and  $\mathbf{b}_{lcl}$  represent the  $n$ th and  $c$ th column of matrix  $\mathbf{G}_{l0}$  and  $\mathbf{B}_{ll}$ , respectively,  $q_{lc}$  is the  $c$ th element of the signal vector  $\mathbf{q}_l$  and  $n_{n0}^{d1}$  is the  $n$ th element of Gaussian noise

vector  $\mathbf{n}_0^{d1}$ . Similarly with the uplink, the ergodic achievable downlink rate of the  $n$ th user in the central cell can be expressed as

$$R_{n0}^{d1} = \mathbb{E} \left\{ \log_2 \left( 1 + \frac{\alpha_0^2 p_{n0}^{d1}}{\sum_{l=1}^L \sum_{c=1}^N p_{cl}^{d1} |\mathbf{g}_{ln0}^T \mathbf{b}_{lcl}|^2 + 1} \right) \right\}. \quad (13)$$

### C. Per-cell Optimization Approach

Based on the achievable rates obtained above, we wish to design a power allocation scheme to optimize the network performance. A global (joint) optimization over all cells would be desirable, but it is found to be both very challenging and not practical due to: 1) high complexity of associated iterative algorithms (based on, e.g., game theoretic ideas), which may have convergence issues as well, and 2) the unbounded and continuous (in space) deployment of BSs which makes it impossible to define an isolated cluster (set of BSs) without being subject to other (out-of-cluster) interference. Therefore, we adopt a more practical but sub-optimal per-cell optimization approach which focuses on a single-cell performance while signals from other cells are regarded as constant interference. More precisely, the sum-rate of a (target) cell is optimized while the interference footprint (caused by surrounding cells) remains unaltered. This approach entails a certain (somewhat simple) level of coordination among cells, according to which power allocation is only performed by a cell when being granted (scheduled) to do so at a particular time slot. A network management entity may, for instance, grant different time slots to different cells according to some scheduling scheme, as exemplified next.

Consider power allocation being applied in an scheduled manner to a set of cells. Cells that do not interfere with each other can perform power allocation simultaneously. Therefore, the entire set of cells can be divided into several groups, where each group consists of geographic distinct cells that do not interfere with each other and can, thus, complete power allocation simultaneously (i.e., similar to the adoption of frequency reuse patterns). Different groups are assigned with different time slots to allocate their transmit powers. Thus, within one scheduling round, each cell has operated with maximized sum rate for at least one time slot. In the remaining slots, there will be a certain rate loss due to the variations in the interference footprint as a result of other cells' optimization. In order to minimize this loss, we will introduce a sum-power constraint in the optimization problem (detailed in the next section), such that the overall interference caused by one cell to others is upper bounded. Moreover, as it will appear, the solution to the optimization problem takes the form of a water filling algorithm.<sup>1</sup> It is then reasonable to expect that most of the power will be often assigned to users in the cell center, which reduces substantially the variations in the interference power as seen by other cells. These observations

<sup>1</sup>The water filling algorithm is a general power assignment algorithm for multichannel systems with a sum-power constraint, which will assign higher powers to channels with better conditions.



are verified by the simulation results in Section V, which shows that this scheduling mechanism can achieve nearly the same performance as the joint (optimal) power allocation after one scheduling round, but with a much lower complexity.

This whole mechanism consists of two procedures: planning the scheduling sequence, and designing an optimal power allocation scheme for each cell. The first problem can be solved by frequency reuse patterns, and the second problem is our focus in this paper.

### III. UPLINK POWER ALLOCATION SCHEMES

Based on the per-cell optimization approach illustrated above, in this section, we aim to find an optimal power allocation scheme to maximize the uplink sum rate in the central cell, which can be obtained by solving the following optimization problem:<sup>2</sup>

$$p_{n0}^{*ul} = \arg \max_{\sum_{c=1}^N p_{c0}^{ul} \leq P^{ul}} \sum_{c=1}^N R_{c0}^{ul}, \quad (14)$$

where  $p_{n0}^{*ul}$  denotes the optimal  $p_{n0}^{ul}$ , and the power allocated to each user is adjusted at the time scale of slow fading. Unfortunately, it is difficult to derive exact solutions for this optimization problem with the expectation in (7). We note that a closed-form expression for the achievable uplink rate with ZF receivers has been given in [26], however, this expression is rather involved and does not readily facilitate the design of an optimized power allocation policy. Instead, we derive tractable lower and upper bounds for the achievable uplink rate, as well as an approximation which lies between these two bounds. Based on these three expressions, we then present their corresponding power allocation strategies obtained as solutions to (14), which are simple and easy to compute. First, we consider the lower bound.

#### A. Lower Bound

##### 1) Closed-form Lower Bound Expression:

*Theorem 1:* The achievable uplink rate of the  $n$ th user in the central cell is lower bounded as

$$R_{n0}^{ul} \geq R_{n0}^{ul,L} = \log_2 \left( 1 + \frac{p_{n0}^{ul} \beta_{0n0} (M - N)}{\sum_{l=1}^L \sum_{c=1}^N p_{cl}^{ul} \beta_{0cl} + 1} \right). \quad (15)$$

*Proof:* See Appendix I.  $\square$

This lower bound is much simpler than the original uplink rate and facilitates the design of the following power allocation strategy.

2) *Power Allocation based on Lower Bound:* With (15), the optimization problem (14) can be simplified by replacing the exact uplink rate  $R_{c0}^{ul}$  with the lower bound  $R_{c0}^{ul,L}$  as follows

$$p_{n0}^{ul} = \arg \max_{\sum_{c=1}^N p_{c0}^{ul} \leq P^{ul}} \sum_{c=1}^N R_{c0}^{ul,L}. \quad (16)$$

<sup>2</sup>As stated in Section II-C, a sum-power constraint (rather than per-user) is imposed to control the interference level.

The following theorem gives the solution to this problem.

*Theorem 2:* The solution to the uplink power allocation problem (16) is<sup>3</sup>

$$p_{n0}^{ul} = \left( \mu_0^{ul} - \frac{1}{d_{n0}} \right)^+, \quad (17)$$

where  $\mu_0^{ul}$  is chosen to satisfy  $\sum_{n=1}^N p_{n0}^{ul} = P^{ul}$  and  $d_{n0} \triangleq \beta_{0n0}(M - N) / \left( \sum_{l=1}^L \sum_{c=1}^N p_{cl}^{ul} \beta_{0cl} + 1 \right)$ .

*Proof:* To maximize the sum rate in (16), it is obvious that the total power needs to be set to the largest value  $P^{ul}$  and each user's power should be nonnegative. Using the Lagrange multiplier approach associated with Karush-Kuhn-Tucker (KKT) conditions, we can obtain the solution to (16), which yields (17).  $\square$

Note that this power allocation is related to the number of BS antennas, the number of users and the large-scale fading coefficients. In particular, as  $M \rightarrow \infty$ ,  $1/d_{n0} \rightarrow 0$ , and the differences in the allocated power among users vanish, meaning that this allocation strategy tends to an equal power assignment with asymptotically large  $M$ . A similar conclusion given in [27] shows that as the signal-to-noise ratio (SNR) grows, the power allocation that maximizes the sum rate converges to an equal power assignment. In our case, this result is obtained by keeping the total SNR fixed, but increasing the number of BS antennas. These observations indicate the equivalence between the two asymptotic regimes, large number of BS antennas and high SNR, which is consistent with the power-scaling results given in [15].

#### B. Upper Bound

##### 1) Closed-form Upper Bound Expression:

*Theorem 3:* The achievable uplink rate of the  $n$ th user in the central cell is upper bounded as

$$R_{n0}^{ul} \leq R_{n0}^{ul,U} = \log_2 \left( 1 + p_{n0}^{ul} \beta_{0n0} (M - N + 1) \times \sum_{h=1}^{\varrho(\mathcal{A}_0)} \sum_{j=1}^{\tau_h(\mathcal{A}_0)} \left\{ \lambda_{h,j}(\mathcal{A}_0) \frac{(-1)^{j-1} \zeta_{0(h)}^{-j}}{(j-1)!} \times \left[ e^{\frac{1}{\zeta_{0(h)}}} \mathbf{E}_1 \left( \frac{1}{\zeta_{0(h)}} \right) - \sum_{m=0}^{j-2} (-1)^m \zeta_{0(h)}^{m+1} m! \right] \right\} \right), \quad (18)$$

where

$$\mathbf{E}_1(x) \triangleq \int_1^\infty \frac{e^{-xt}}{t} dt, \quad \text{Re}(x) \geq 0, \quad (19)$$

is the exponential integral function of order 1,  $\zeta_{0k} \triangleq p_{cl}^{ul} \beta_{0cl}$  ( $l = 1, \dots, L; c = 1, \dots, N$ ) with  $k = N(l-1) + c$ . Moreover,  $\mathcal{A}_0 \triangleq \text{diag}(\zeta_{01}, \zeta_{02}, \dots, \zeta_{0T})$ , while  $T = NL$ ,  $\varrho(\mathcal{A}_0)$  is the number of the distinct diagonal elements of  $\mathcal{A}_0$ ,  $\zeta_{0\langle 1 \rangle} > \zeta_{0\langle 2 \rangle} > \dots > \zeta_{0\langle \varrho(\mathcal{A}_0) \rangle}$  are the distinct diagonal elements in decreasing order,  $\tau_h(\mathcal{A}_0)$  is the multiplicity of  $\zeta_{0(h)}$ , and  $\lambda_{h,j}(\mathcal{A}_0)$  is the  $(h, j)$ th characteristic coefficient of  $\mathcal{A}_0$  [28].

<sup>3</sup>Observe the water-filling form of the solution. On the uplink, users get their transmit power information through the feedback from the BS, which acquires the channel information (large-scale fading) for all users and determines the uplink transmit powers. Here, the user scheduling involving fairness is not considered, but will be considered for future work.

*Proof:* See Appendix II.  $\square$

We can now use this upper bound to provide a corresponding power allocation strategy.

2) *Power Allocation based on Upper Bound:* After replacing the exact uplink rate  $R_{c0}^{ul}$  in (14) with the upper bound  $R_{c0}^{ul,U}$  given in (18), the central cell optimization problem becomes

$$p_{n0}^{ul} = \arg \max_{\sum_{c=1}^N p_{c0}^{ul} \leq P^{ul}} \sum_{c=1}^N R_{c0}^{ul,U}, \quad (20)$$

and the solution is presented in the following theorem.

*Theorem 4:* The solution to the uplink power allocation problem (20) is

$$p_{n0}^{ul} = \left( \mu_0^{ul} - \frac{1}{k_{n0}} \right)^+, \quad (21)$$

where  $\mu_0^{ul}$  is chosen to satisfy  $\sum_{n=1}^N p_{n0}^{ul} = P^{ul}$  and

$$k_{n0} \triangleq \beta_{0n0}(M - N + 1) \sum_{h=1}^{\varrho(\mathcal{A}_0)} \sum_{j=1}^{\tau_h(\mathcal{A}_0)} \left\{ \lambda_{h,j}(\mathcal{A}_0) \frac{(-1)^{j-1} \zeta_{0(h)}^{-j}}{(j-1)!} \right. \\ \left. \times \left[ e^{\frac{1}{\zeta_{0(h)}}} \mathbf{E}_1 \left( \frac{1}{\zeta_{0(h)}} \right) - \sum_{m=0}^{j-2} (-1)^m \zeta_{0(h)}^{m+1} m! \right] \right\}. \quad (22)$$

*Proof:* Follows the same steps as the proof of *Theorem 2*.

Observe that, despite the involved structure of (22), the coefficients  $k_{n0}$  (for different users) differ only through the factor  $\beta_{0n0}$  and the remaining factors can be regarded as constant. Moreover, as  $M \rightarrow \infty$ ,  $1/k_{n0} \rightarrow 0$ , and the differences among users in this power allocation scheme disappear. Again, a tendency towards an equal assignment policy is observed with  $M$  growing without limit, which agrees with the observation in *Theorem 2*.

Next, we give a third tractable expression for the achievable uplink rate—a new approximation, which is proven to lie between the upper and lower bounds derived above. As shown later in this paper, our simulations demonstrate that this result is particularly accurate over a wide range of operating conditions.

### C. Approximation

1) *Closed-form Approximation:* To obtain the approximation, a key useful tool is given first. To the best of our knowledge, this approximation is new and provides a useful general tool for studying ergodic capacity.

*Lemma 1:* If  $X$  and  $Y$  are independent positive random variables, then

$$\mathbb{E} \left\{ \log_2 \left( 1 + \frac{X}{Y} \right) \right\} \approx \log_2 \left( 1 + \frac{\mathbb{E}\{X\}}{\mathbb{E}\{Y\}} \right), \quad (23)$$

where

$$\log_2 \left( 1 + \frac{1}{\mathbb{E}\left\{ \frac{Y}{X} \right\}} \right) \leq \log_2 \left( 1 + \frac{\mathbb{E}\{X\}}{\mathbb{E}\{Y\}} \right) \leq \log_2 \left( 1 + \mathbb{E} \left\{ \frac{X}{Y} \right\} \right). \quad (24)$$

*Proof:* See Appendix III.  $\square$

Note that the approximation in this lemma lies between two bounds obtained by a direct application of Jensen's

inequality. Therefore, we can conclude that  $\log_2 \left( 1 + \frac{\mathbb{E}\{X\}}{\mathbb{E}\{Y\}} \right)$  is a relatively tight approximation of  $\mathbb{E} \left\{ \log_2 \left( 1 + \frac{X}{Y} \right) \right\}$  (at least tighter than one of the two bounds). It is also important to note that the lower and upper bounds given respectively in *Theorems 1* and *3* are obtained by applying the corresponding bounds in (24). Moreover, this approximation becomes asymptotically exact as shown in the following corollary.

*Corollary 1:* Let  $X$  and  $Y$  in *Lemma 1* be given as

$$X = 1 / \sum_{i=1}^{\theta_1} x_i, \quad Y = \sum_{j=1}^{\theta_2} y_j, \quad (25)$$

where  $x_i$  ( $i = 1, \dots, \theta_1$ ) and  $y_j$  ( $j = 1, \dots, \theta_2$ ) are both sequences of positive square-integrable random variables which are not necessarily independent across  $i$  and  $j$ , respectively, and as  $\theta_1 \rightarrow \infty$ ,  $\frac{\theta_2}{\theta_1} \rightarrow \omega \in (1, \infty)$ . Moreover, there exist  $\epsilon_1 > 0$  and  $\epsilon_2 > 0$ , such that

$$\liminf_i \mathbb{E} \{x_i\} > \epsilon_1, \quad \liminf_j \mathbb{E} \{y_j\} > \epsilon_2. \quad (26)$$

Then, as  $\theta_1 \rightarrow \infty$ ,

$$\mathbb{E} \left\{ \log_2 \left( 1 + \frac{X}{Y} \right) \right\} - \log_2 \left( 1 + \frac{\mathbb{E}\{X\}}{\mathbb{E}\{Y\}} \right) \xrightarrow{a.s.} 0. \quad (27)$$

*Proof:* See Appendix IV.  $\square$

When *Lemma 1* is applied to the achievable rate analysis,  $\theta_1$  denotes the number of BS antennas. Therefore, according to *Corollary 1*, the approximation (23) will be tighter as the number of BS antennas increases and becomes asymptotically exact as  $M \rightarrow \infty$ . Note that a similar corollary can be easily obtained to show the asymptotic exactness of (23) for the case  $X = \sum_{i=1}^{\theta_1} x_i$ , which becomes useful when analyzing maximum ratio combining receivers.

Based on these results, the following theorem gives a tractable expression for the achievable uplink rate in (7).

*Theorem 5:* The achievable uplink rate for the  $n$ th user in the central cell is approximated as

$$R_{n0}^{ul} \approx \tilde{R}_{n0}^{ul} = \log_2 \left( 1 + \frac{p_{n0}^{ul} \beta_{0n0} (M - N + 1)}{\sum_{l=1}^L \sum_{c=1}^N p_{cl}^{ul} \beta_{0cl} + 1} \right). \quad (28)$$

*Proof:* Divide the numerator and denominator of (7) by  $\|\mathbf{a}_{0n0}\|^2$ . Then,

$$R_{n0}^{ul} = \mathbb{E} \left\{ \log_2 \left( 1 + \frac{p_{n0}^{ul} \beta_{0n0} z_{0n0}}{\sum_{l=1}^L \sum_{c=1}^N p_{cl}^{ul} |\xi_{0cl}|^2 + 1} \right) \right\}. \quad (29)$$

From the proof of *Theorem 1*, we know that  $z_{0n0}$  and  $\xi_{0cl}$  are independent. Hence, applying *Lemma 1* in (29) yields

$$R_{n0}^{ul} \approx \tilde{R}_{n0}^{ul} = \log_2 \left( 1 + \frac{p_{n0}^{ul} \beta_{0n0} \mathbb{E} \{z_{0n0}\}}{\sum_{l=1}^L \sum_{c=1}^N p_{cl}^{ul} \mathbb{E} \{|\xi_{0cl}|^2\} + 1} \right). \quad (30)$$

Plugging (40) and (46) into (30) leads to the desired result.  $\square$

Note from *Lemma 1* that the approximation in (28) lies between the lower and upper bounds in (15) and (18), and as

shown later in the numerical discussion, this approximation is more accurate than the aforementioned bounds. Moreover, according to *Corollary 1*, this approximation becomes progressively more accurate as the number of BS antennas increases. Next, we will use it for the power allocation analysis.

2) *Power Allocation based on Approximation:* With the tractable approximation (28), the central cell optimization problem becomes

$$p_{n0}^{u1} = \arg \max_{\sum_{c=1}^N p_{c0}^{u1} \leq P^{u1}} \sum_{c=1}^N \tilde{R}_{c0}^{u1}, \quad (31)$$

and the following theorem presents the solution.

*Theorem 6:* The solution to the uplink power allocation problem (31) is

$$p_{n0}^{u1} = \left( \mu_0^{u1} - \frac{1}{t_{n0}} \right)^+, \quad (32)$$

where  $\mu_0^{u1}$  satisfies  $\sum_{n=1}^N p_{n0}^{u1} = P^{u1}$  and  $t_{n0} \triangleq \beta_{0n0}(M - N + 1) / \left( \sum_{l=1}^L \sum_{c=1}^N p_{cl}^{u1} \beta_{0cl} + 1 \right)$ .

*Proof:* Follows the same steps as the proof of *Theorem 2*.  $\square$

When  $M \gg N$ , the approximation in (28) approaches the lower bound in (15) asymptotically (as  $M - N \rightarrow \infty$ ). Therefore, for  $M \gg N$ , the power allocation scheme based on the approximation is also equivalent to that based on the lower bound. Hence, the power allocation strategy in *Theorem 6* also tends to an equal power assignment as  $M \rightarrow \infty$ .

We have provided three different power allocation schemes based on three tractable expressions for the achievable rate: an upper bound, a lower bound, and an approximation which lies between these two bounds. In Section V, we will evaluate the performance of these three schemes with some numerical comparisons under different operating conditions.

#### IV. DOWNLINK POWER ALLOCATION SCHEME

Similarly with the uplink case, we assume the total power transmitted by one BS is constrained as  $\sum_{c=1}^N p_{c0}^{d1} \leq P^{d1}$ . Then, we aim to find an optimal power assignment scheme to maximize the cell sum rate. That is, we need to solve the optimization problem:

$$p_{n0}^{d1} = \arg \max_{\sum_{c=1}^N p_{c0}^{d1} \leq P^{d1}} \sum_{c=1}^N R_{c0}^{d1}, \quad (33)$$

where  $p_{n0}^{d1}$  denotes the optimal  $p_{n0}^{d1}$ , and the power allocated to each user is adjusted at the time scale of slow fading. The exact solution of this optimal problem is hard to obtain. Therefore, aiming to obtain a simple power allocation scheme, we first present a new lower bound for the achievable downlink rate.

##### A. Lower Bound

We first require the following preliminary result.

*Lemma 2:* To satisfy the constraint (9), the constant scalar  $\alpha_l$  for the ZF precoding matrix  $\mathbf{B}_{ll}$  is given by

$$\alpha_l = \sqrt{\frac{M - N}{\sum_{n=1}^N 1/\beta_{lnl}}}. \quad (34)$$

*Proof:* See Appendix V.  $\square$

We now give a tractable lower bound for the achievable downlink rate.

*Theorem 7:* The achievable downlink rate of the  $n$ th user in the central cell is lower bounded by

$$R_{n0}^{d1} \geq R_{n0}^{d1,L} = \log_2 \left( 1 + \frac{p_{n0}^{d1}(M - N)/\Lambda_{00}}{\sum_{l=1}^L \sum_{c=1}^N \frac{p_{cl}^{d1} \beta_{ln0}}{\beta_{lcl} \Lambda_{ll}} + 1} \right), \quad (35)$$

where  $\Lambda_{ll} \triangleq \sum_{k=1}^N 1/\beta_{lkl}$ .

*Proof:* See Appendix VI.  $\square$

Note that this lower bound is more tractable than the exact downlink rate (13). Since the ZF precoding vector  $\mathbf{b}_{lcl}$  in (13) is different across cells and users, the approach to obtain the upper bound and approximation in the uplink case does not apply here. Therefore, in the following subsection, we use this lower bound to conduct the power allocation analysis.

##### B. Power Allocation

With (35), the optimization problem (33) can be simplified by replacing the exact uplink rate  $R_{c0}^{d1}$  with the lower bound  $R_{c0}^{d1,L}$  as

$$p_{n0}^{d1} = \arg \max_{\sum_{c=1}^N p_{c0}^{d1} \leq P^{d1}} \sum_{c=1}^N R_{c0}^{d1,L}, \quad (36)$$

and the solution is given by the following theorem.

*Theorem 8:* The solution to the downlink power allocation problem (36) is<sup>4</sup>

$$p_{n0}^{d1} = \left( \mu_0^{d1} - \frac{1}{s_{n0}} \right)^+, \quad (37)$$

where  $\mu_0^{d1}$  satisfies  $\sum_{n=1}^N p_{n0}^{d1} = P^{d1}$  and  $s_{n0} \triangleq ((M - N)/\Lambda_{00}) / \left( \sum_{l=1}^L \sum_{c=1}^N \frac{p_{cl}^{d1} \beta_{ln0}}{\beta_{lcl} \Lambda_{ll}} + 1 \right)$ .

*Proof:* The derivation follows the same procedure as that shown in *Theorem 2*.  $\square$

Note that the effective power allocation for the downlink is related to the number of BS antennas, the number of users and the large-scale fading coefficient. Similarly with the uplink, as  $M \rightarrow \infty$ ,  $1/s_{n0} \rightarrow 0$ . Hence, the allocated power  $p_{n0}^{d1}$  will approach  $P^{d1}/N$ , that is, the equal power assignment. Therefore, an important insight we can draw from comparing *Theorem 4* and *6* is that regardless of uplink or downlink, the equal power allocation policy tends to be optimal as the number of BS antennas grows without limit. That means that large antenna arrays produce the same effect as increasing the transmit power and thus they can be used to cut down the transmit power for both uplink and downlink.

<sup>4</sup>The channel knowledge needed in the downlink power allocation (large-scale fading coefficient) is acquired from the uplink due to channel reciprocity.

## V. NUMERICAL RESULTS

In this section, we provide numerical results for a set of 19 cells with radius  $r_c = 1000$  meters in Fig. 1.  $N$  users are distributed randomly and uniformly in each cell, with the exclusion of a central disk of radius  $r_h = 100$  meters around the BSs. The large scale fading is modeled using  $\beta_{icl} = z_{icl}/(r_{icl}/r_h)^v$ , where  $z_{icl}$  is a log-normal random variable with standard deviation  $\sigma$ ,  $v$  is the path loss exponent, and  $r_{icl}$  is the distance between the  $c$ th user in the  $l$ th cell and the  $i$ th BS. In our simulations, we choose  $\sigma = 8$  dB and  $v = 3.8$ . The small-scale fading is assumed to be Rayleigh distributed. Here, we consider a simple case where only adjacent cells produce interference (while further cells are neglected). Then, according to the scheduling approach described in Section II-C, the whole set of cells can be divided into three groups, marked as 1, 2 and 3 in Fig. 1, and each group conducts power allocation in a single time slot. Since this scheduling approach is based on the per-cell optimized powers, we will first focus on the performance of a single target cell (central cell in Fig. 1), during the first time slot in which interference conditions (i.e., the powers in the other interfering cells labeled 2 and 3) are fixed. The network performance, taking into account dynamical effects of the interference footprint, will be considered subsequently, where we evaluate the sum rate of the entire 19 cells and compare our scheduled power allocation with the joint (optimal) power allocation.

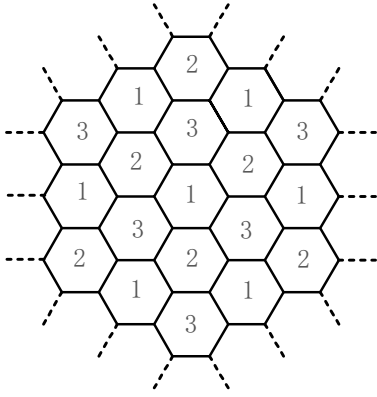


Fig. 1: Simulation model.

### A. One-cell Performance

1) *Uplink*: For the uplink, we first evaluate the accuracy of the three expressions for the achievable rate and the performance of the power allocation schemes based on them. Then, the best power allocation strategy is selected to quantify the benefits over equal power assignment in a range of scenarios with variable  $M/N$  ratios. As we will see, the approximation is the most accurate expression and the corresponding power allocation scheme can gain noticeable improvements over the equal power assignment for practical  $M/N$  ratios. The transmit powers of users in surrounding 6 cells are set to 10 dB for the uplink analysis.

In Fig. 2, the simulated uplink sum rate in (7) is compared with its lower bound in (15), upper bound in (18) and approximation in (28). Results are given under two different total transmit powers—20 and 30 dB, and the transmit power of

each user is assumed to be the same. Clearly, in all cases, the uplink rate grows with the increasing number of BS antennas. Moreover, we see a close agreement between the simulation results and our analytical lower bound and approximation. In particular, the approximation is almost indistinguishable with the simulated values and lies between the lower and upper bound, as expected from Section III-C. Observe that the lower bound and the approximation are very close to each other; indeed, this can be anticipated from (15) and (28) when  $M \gg N$ .

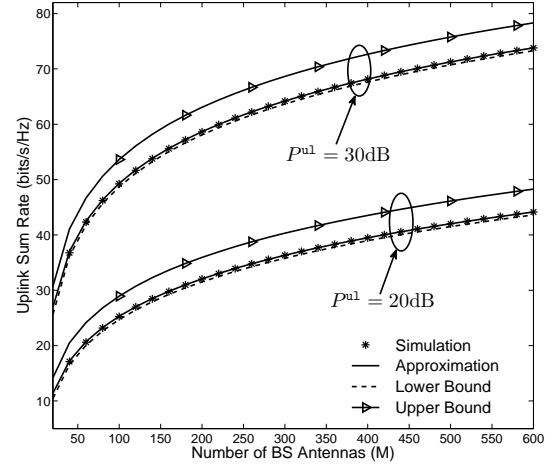


Fig. 2: Uplink sum rate per cell vs. the number of BS antennas with equal power allocation, where  $N = 10$  users.

To show the distinction between the lower bound and approximation, we present simulation curves for  $M$  closer to  $N$  in Fig. 3. Here, the simulated uplink sum rate as well as the bounds and approximation are plotted against the uplink total transmit power. We can see that, with  $M$  closer to  $N$ , the gap between the lower bound and approximation becomes more evident than that in Fig. 2, and in this case, the accuracy of the upper bound is comparable to that of the lower bound. Importantly, the approximation is still the most accurate. Next, we investigate the proposed power allocation schemes.

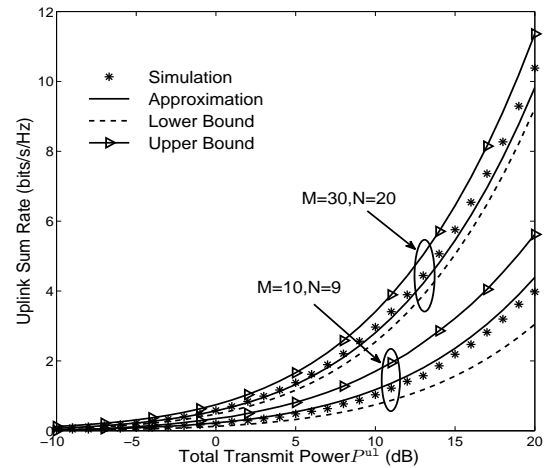


Fig. 3: Uplink sum rate per cell vs. the total transmit power  $P^ul$ , with equal power allocation.

Fig. 4 plots the different power allocation schemes described



in Section III under both multicell and single-cell scenarios.<sup>5</sup> For multicell systems, the power allocation (17) based on the lower bound, (21) based on the upper bound and (32) based on the approximation are compared with the equal power assignment. Clearly, all power allocation schemes achieve noticeable improvements on the uplink rate, especially when the number of BS antennas is within a few hundred. As expected, these improvements become less significant as the number of BS antennas increases, and we observe that, when  $M \rightarrow \infty$ , the proposed power allocation schemes approach equal power assignments. It can be observed that the power allocation based on the approximation leads to the maximum enhancement, with a very slight improvement over the assignment based on the lower bound. This is a direct consequence of the increased accuracy of the approximation over the bounds.

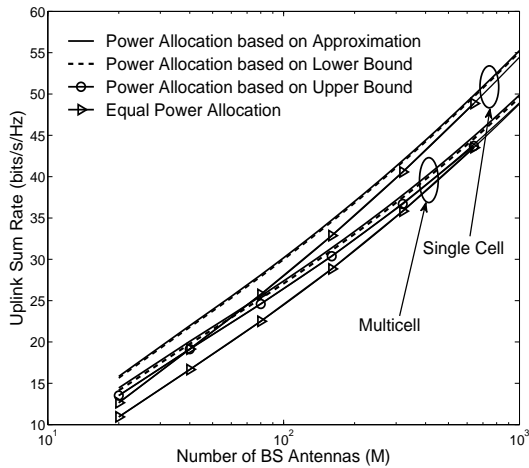


Fig. 4: Uplink sum rate vs. the number of BS antennas with different power allocation schemes, where  $P^{ul} = 20\text{dB}$  and  $N = 10$  users.

To illustrate the effect of these power allocation schemes further, we define the *relative gain* of the power allocation as  $\eta \triangleq (C_{PA} - C_{EQ})/C_{EQ}$ , where  $C_{PA}$  denotes the sum rate with our power allocation schemes, and  $C_{EQ}$  denotes the sum rate under equal power allocation.

Fig. 5 shows that, for the proposed power allocation schemes, significant gains are attained even with a few hundred antennas. For example, when  $M = 100$ , the power allocation based on the approximation in multicell systems can obtain almost 14% gain. Moreover, the gap in the multicell scenario is enhanced compared with that in the single-cell scenario. This is due to the addition of interference from other cells, which makes each user's receive SNR smaller, thus leading to a better performance of the water filling algorithm. We thus anticipate this difference to grow as the number of cells increases.

The above results indicate that the gain brought by the power allocation schemes diminishes as the number of antennas increases while keeping the number of users fixed, i.e., as  $M/N \rightarrow \infty$ , which is sometimes referred to as the “massive MIMO regime”. However, this might not be representative of practical scenarios where (i) the number of

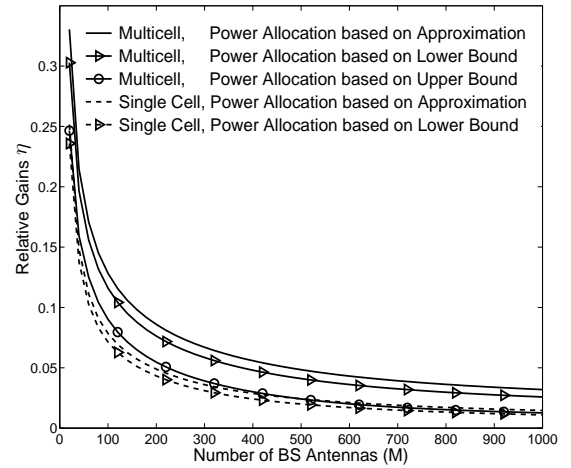


Fig. 5: Relative gains of different uplink power allocation schemes, where  $P^{ul} = 20\text{dB}$  and  $N = 10$ .

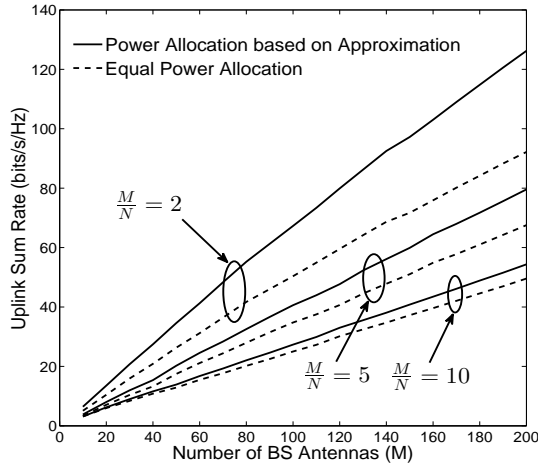
antennas is limited or (ii) the number of users is also large and is comparable with the number of antennas. In order to explore the effects of the proposed schemes in such scenarios, we next focus on the regime where  $M$  and  $N$  scale together with a fixed  $M/N$  ratio. Since the power allocation based on the analytical approximation (32) has been shown to attain the maximum gain, we only consider this strategy in the subsequent simulations.

Fig. 6 shows the effect of optimized power allocation with different  $M/N$  ratios. We find that under a fixed  $M/N$ , the gain brought by the power allocation grows as the number of BS antennas increases, which differs from the conclusion under a fixed  $N$ . Since  $M$  and  $N$  scale together in this case, a larger number of antennas implies a larger number of users and, consequently, an increased interference which makes the receive SNR smaller. Notably, it is precisely under low SNR conditions where the power allocation optimization brings most benefit. The same arguments explain the fact that larger gains are attained at smaller  $M/N$ . It is also interesting to note that the curves in Fig. 6 are all approximately linear with no offset, which suggests that the relative rate gap between schemes with and without power allocation remains constant no matter the value of  $M$ . Therefore, the relative gain for each  $M/N$  ratio can be obtained from the slopes of these curves via linear fitting. These results are shown in Fig. 7.

Fig. 7 presents how the relative gains of the power allocation change with the  $M/N$  ratio. We give results for four different total transmit powers—10, 15, 20 and 25dB. As expected, the relative gains reduce with increasing  $M/N$  or  $P^{ul}$ , which is consistent with the property of the water filling algorithm. Importantly, Fig. 7 allows determining the  $M/N$  ratios which bring gains over a particular (required) threshold, and these results can be used to decide whether the power allocation is justified. For example, if we set the tolerable minimum relative gain as 10%, the  $M/N$  ratio with transmit power 10 and 15dB can be high; for example, more than 20. This implies that even if there are more than 20 BS antennas per user, one still obtains more than 10% relative gain by using optimized power allocation. However, as the transmit power is boosted, the relative gains of power allocation will diminish. For example,

<sup>5</sup>The power allocation schemes in single-cell scenario are obtained by setting  $L = 0$ . Then, the power allocation based on the approximation and that based on the upper bound have the same expression.



Fig. 6: Uplink sum rate vs. the number of BS antennas with constant  $M/N$ .

for a transmit power of 20dB, relative gains exceeding 10% are only observed for 12 BS antennas or less per user (i.e.,  $M/N \leq 12$ ); whilst for a transmit power 25dB, these gains are achieved for only 4 BS antennas or less per user. For system configurations in which the relative gain does not exceed a prescribed threshold (e.g., 10%), one may favor equal power allocation, as a consequence of its minimal complexity. This is further explored in Table I, which indicates the maximum number of BS antennas per user (i.e., max value of  $M/N$ ) required for two prescribed relative gain thresholds (10% and 20%), for different transmit power constraints. These results demonstrate that, even when there are 10 times more BS antennas than users, and even for moderate transmit powers (e.g., 15dB), one can still obtain quite substantial gains (e.g., more than 20%) by performing power allocation optimization.

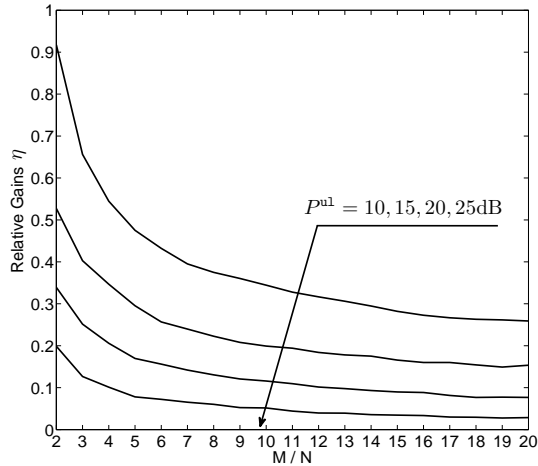


Fig. 7: Relative gains of the power allocation based on approximation.

TABLE I: Maximum  $M/N$  ratios with different uplink relative gain thresholds and total transmit powers

Minimum Relative Gain	$P^{ul}$			
	10dB	15dB	20dB	25dB
10%	93	30	12	4
20%	29	10	4	2

2) *Downlink*: For the downlink, we first show the performance of the lower bound for the achievable rate. Then, the corresponding power allocation scheme is evaluated for two different kinds of users—central users and edge users. We find that although the power allocation strategy still tends to an equal power assignment when  $M/N \rightarrow \infty$ , it obtains considerable gains with practical  $M/N$  values. The transmit powers in surrounding 6 cells are set to 30dB for the downlink analysis.

In Fig. 8, the simulated downlink sum rate in (13) is compared with the analytical lower bound in (35) under two different total transmit powers—40 and 50dB. The transmit power intended for each user is assumed to be the same. Clearly, in all cases, the downlink sum rates grow as the number of BS antennas increases and we can see a close agreement between the simulation results and our analytical lower bound.

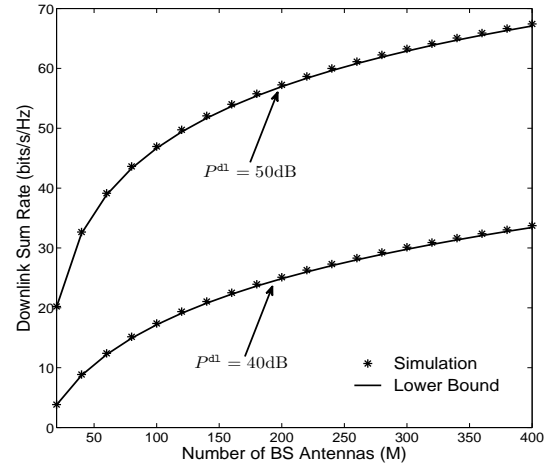
Fig. 8: Downlink sum rate per cell vs. the number of BS antennas with equal power allocation, where  $N = 10$  users.

Fig. 9 shows the effect of downlink power allocation based on the lower bound (37). Results are given for central and edge users, which are separated by  $0.8r_c$ . The power allocation scheme has a noticeable rate improvement for edge users, but for central users the improvement is almost negligible. This is because the SNR of edge users is much smaller than that of central users and, as discussed above, the gains due to power allocation are most prominent at low SNRs. In addition, when the number of BS antennas increases, the improvement brought by power allocation also reduces, and as  $M \rightarrow \infty$ , this power allocation scheme tends to an equal power allocation. These observations are all in accordance with our analytical results.

To further illustrate the improvement due to power allocation for the downlink, Fig. 10 shows the relative gains of the proposed power allocation scheme over the equal power assignment. This power allocation policy provides the maximum enhancement for edge users and, as anticipated, these gains are almost negligible for central users. As expected, the gains diminish with  $M$ , vanishing asymptotically as  $M \rightarrow \infty$ , which means that the equal power allocation tends to be optimal. Note however that the gains for edge users are still noticeable for antenna numbers up to a few hundred. For example, taking

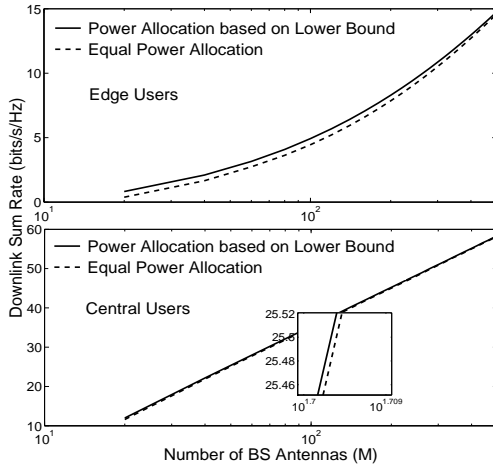


Fig. 9: Downlink sum rate of central and edge users vs. the number of BS antennas with the proposed power allocation scheme, where  $P^{d1} = 40\text{dB}$  and  $N = 10$  users.

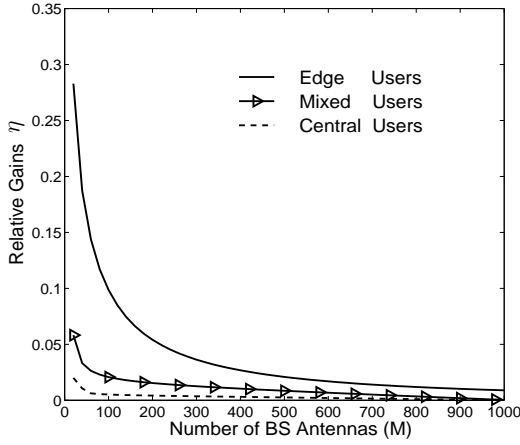


Fig. 10: Relative gains of the downlink power allocation scheme, where  $P^{d1} = 40\text{dB}$  and  $N = 10$  users.

$M = 100$ , power allocation yields a 10% gain on the downlink rate for edge users, in contrast to a 2% gain for central users.

Similarly to the uplink analysis, for the downlink, we are also interested in further exploring when the power allocation scheme is justified. Due to the precoding matrix constraint, the downlink sum rate does not show the linear behavior shown in Fig. 6 for the uplink (meaning that the sum rate depends on  $M$ , and not only on the ratio  $M/N$ ). Therefore, we prescribe some specific values of  $M$  and  $N$  to perform this analysis, where we only consider edge users. Table II.a presents the maximum numbers of antennas for a minimum acceptable relative gain with different numbers of users and transmit powers. We can see that, as expected, when the total transmit power or the relative gain threshold increase, the maximum  $M$  decreases. Moreover, the maximum  $M$  grows as  $N$  increases and this is because more users result in more interference and lower SNR, which ultimately leads to a higher relative gain from the power allocation. In Table II.b, we present the minimum numbers of users for an acceptable relative gain with different  $M$  and transmit powers. These two tables can be used to determine whether the downlink power allocation is justified. For example, when  $N = 10$  with transmit power 35dB, the power allocation guarantees a minimum relative gain of 10%

for any  $M$  up to 288. But for 40 and 45dB, such gains are achieved for  $M$  up to 99 and 38, respectively. Similarly, if  $M$  is set to 100, according to Table II.b, a system with transmit power 35dB can obtain at least 10% relative gain with more than 7 users. But for 40 and 45dB, to achieve this target, there must be more than 10 and 16 users, respectively. As for systems that can not reach the prescribed minimum relative gain, the equal power allocation may be more suitable.

TABLE II: Relationship between  $M$  and  $N$  with different downlink relative gain thresholds and total transmit powers

$N$	Minimum Relative Gain	$P^{d1}$		
		35dB	40dB	45dB
5	10%	48	18	9
	20%	16	8	6
10	10%	288	99	38
	20%	89	36	18
20	10%	1442	463	159
	20%	454	147	63

(a) Maximum  $M$  with different  $N$

$M$	Minimum Relative Gain	$P^{d1}$		
		35dB	40dB	45dB
50	10%	5	7	12
	20%	8	12	18
100	10%	7	10	16
	20%	11	17	25
200	10%	9	14	22
	20%	14	25	36

(b) Minimum  $N$  with different  $M$

## B. Multicell Performance with Scheduling Mechanism

The previous results demonstrated the performance from the perspective of a given cell, assuming that the interference power originating from the surrounding cells was fixed. This allowed us to focus on the relative pros and cons of the different power allocation strategies under ranging conditions. Now, we adopt a network perspective, and evaluate the performance of a set of cells with our scheduling mechanism for power allocation, taking the uplink as an example. We adopt the power allocation strategy based on the analytical approximation (32), as it was shown in Section V-A.1 to yield the best performance among the different schemes considered. We apply this strategy for each cell and evaluate the sum rate of the 19 cells in Fig. 1.<sup>6</sup> The initial transmit power for each user is 10dB. In time slot  $i$  (for  $i = 1, 2, 3, \dots$ ), only cells within group  $j = \text{mod}(i-1, 3) + 1$  adjust their powers. Thus, during each scheduling round (comprising 3 time slots), each cell will operate with optimized power for one slot. In the subsequent two time slots, it will incur some performance loss due to a changing interference footprint caused by power allocation being performed in the surrounding cells. As a benchmark, we compare with the optimal performance achieved via joint

<sup>6</sup>When evaluating the performance of these 19 cells, some adjacent cells out of this cluster, which bring interference to cells at the edge of this cluster, should also be considered. Since the number of power allocation objectives must be finite, we assume these out-of-cluster cells have constant transmit powers.

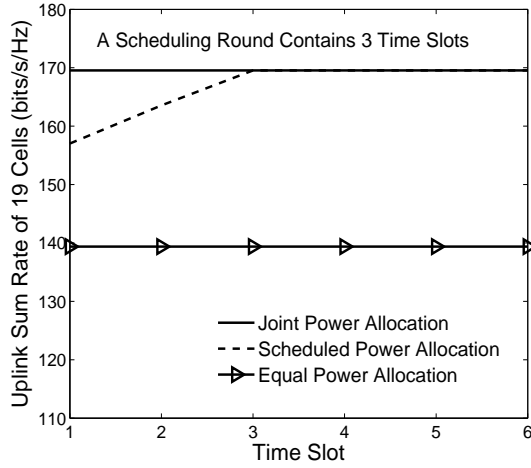


Fig. 11: The uplink sum rate of 19 cells vs. the index of time slots, where  $M = 20$ ,  $N = 5$  and  $P^{ul} = 50W$ .

power allocation across all 19 cells,<sup>7</sup> and with the simple equal power allocation strategy (i.e., no power adaptation).

From Fig. 11, we find that, in slots 1 and 2, the joint method performs better than the scheduling one. But from slot 3, the gap between them is negligible. That is, after a single scheduling round, the scheduled power allocation achieves almost the same performance as the joint one, and this nearly optimal performance is maintained in the subsequent time slots. The joint method, highly complex and requiring high-level communication among BSs, entails a challenging implementation in practice. At a much lower complexity, the scheduled power allocation achieves nearly the same result (with differences inappreciable) after the first scheduling round which can be seen as a transient effect. Therefore, considering both complexity and performance, the scheduling mechanism is an attractive solution.

## VI. CONCLUSION

We have characterized the gains in achievable rate brought by power allocation schemes in multicell MIMO systems with large antenna arrays. With ZF receivers, we have derived lower and upper bounds for the achievable uplink rate and an approximation which lies between these two bounds. With ZF precoders, we have derived a relatively simple lower bound for the achievable downlink rate. As opposed to a joint (across all cells) optimization, we have proposed a much simpler scheduling method to plan the power allocation arrangements for different cells. Based on this, new power allocation strategies have been presented, which can bring considerable gains over the equal power allocation assignment. These gains are substantial for moderate array sizes up to a few hundred, and for the case where  $N$  is allowed to scale with  $M$ , these gains increase as  $M$  grows with fixed  $M/N$ . Based on these results, we provide applicable values of  $M/N$  with a certain minimum relative gain in the achievable rates, which can be used as practical design rules to justify the use of optimized

<sup>7</sup>The joint power allocation problem does not admit an analytical solution; thus, it was evaluated numerically using the function “fmincon” in the MATLAB optimization toolbox.

power allocation schemes over the equal assignment policy. This reveals the applicability of our allocation schemes under a wide range of scenarios with practical numbers of users and antennas. From a network point of view, the proposed scheduling mechanism achieves almost the same performance as the joint power allocation after one scheduling round, with much reduced complexity.

## APPENDIX I PROOF OF THEOREM 1

Application of Jensen’s inequality to (7) gives

$$R_{n0}^{ul} \geq R_{n0}^{ul,L} = \log_2 \left( 1 + \frac{p_{n0}^{ul}}{\sum_{l=1}^L \sum_{c=1}^N p_{cl}^{ul} \mathbb{E} \{ |\mathbf{a}_{0n0}^H \mathbf{g}_{0cl}|^2 \} + \mathbb{E} \{ \|\mathbf{a}_{0n0}\|^2 \}} \right), \quad (38)$$

which can be further written as

$$R_{n0}^{ul,L} = \log_2 \left( 1 + \frac{p_{n0}^{ul}}{\sum_{l=1}^L \sum_{c=1}^N p_{cl}^{ul} \mathbb{E} \{ |\xi_{0cl}|^2 \|\mathbf{a}_{0n0}\|^2 \} + \mathbb{E} \{ \|\mathbf{a}_{0n0}\|^2 \}} \right), \quad (39)$$

where  $\xi_{0cl} \triangleq \mathbf{a}_{0n0}^H \mathbf{g}_{0cl} / \|\mathbf{a}_{0n0}\|$ . Since  $\mathbf{g}_{0cl}$  has a rotation-invariant distribution and  $\frac{\mathbf{a}_{0n0}^H}{\|\mathbf{a}_{0n0}\|}$  can be regarded as a column of a rotation matrix,  $\xi_{0cl}$  has the same distribution as the elements of  $\mathbf{g}_{0cl}$  and is independent of  $\mathbf{a}_{0n0}$ . Therefore,

$$\mathbb{E} \{ |\xi_{0cl}|^2 \} = \beta_{0cl}, \quad (40)$$

and (39) can be expressed as

$$R_{n0}^{ul,L} = \log_2 \left( 1 + \frac{p_{n0}^{ul}}{\left( \sum_{l=1}^L \sum_{c=1}^N p_{cl}^{ul} \mathbb{E} \{ |\xi_{0cl}|^2 \} + 1 \right) \mathbb{E} \{ \|\mathbf{a}_{0n0}\|^2 \}} \right). \quad (41)$$

Let  $z_{0n0} \triangleq 1 / [(\mathbf{H}_{00}^H \mathbf{H}_{00})^{-1}]_{nn}$ , which is chi-squared distributed with probability density [29]

$$f(z_{0n0}) = \frac{e^{-z_{0n0}}}{\Gamma(M-N+1)} (z_{0n0})^{M-N}, \quad z_{0n0} \geq 0, \quad (42)$$

and therefore,  $\mathbb{E} \{ 1/z_{0n0} \} = 1/(M-N)$ . Then, with  $\|\mathbf{a}_{0n0}\|^2 = [(\mathbf{G}_{00}^H \mathbf{G}_{00})^{-1}]_{nn} = 1/\beta_{0n0} z_{0n0}$ ,

$$\mathbb{E} \{ \|\mathbf{a}_{0n0}\|^2 \} = 1/\beta_{0n0} (M-N). \quad (43)$$

Plugging (40) and (43) into (41) yields the result.

## APPENDIX II PROOF OF THEOREM 2

Applying Jensen’s inequality in (7) gives

$$R_{n0}^{ul} \leq R_{n0}^{ul,U} = \log_2 \left( 1 + \mathbb{E} \left\{ \frac{p_{n0}^{ul}}{\sum_{l=1}^L \sum_{c=1}^N p_{cl}^{ul} |\mathbf{a}_{0n0}^H \mathbf{g}_{0cl}|^2 + \|\mathbf{a}_{0n0}\|^2} \right\} \right). \quad (44)$$

Recalling the definition of  $\xi_{0cl}$  and  $z_{0n0}$  in the proof of *Theorem 1*, and the fact that  $\xi_{0cl}$  is independent of  $z_{0n0}$ , we obtain

$$R_{n0}^{u1,u} = \log_2 \left( 1 + \mathbb{E} \{ p_{n0}^{u1} \beta_{0n0} z_{0n0} \} \mathbb{E} \left\{ \frac{1}{\sum_{l=1}^L \sum_{c=1}^N p_{cl}^{u1} |\xi_{0cl}|^2 + 1} \right\} \right). \quad (45)$$

Recalling (42), we can get the following expectation

$$\mathbb{E} \{ p_{n0}^{u1} \beta_{0n0} z_{0n0} \} = p_{n0}^{u1} \beta_{0n0} (M - N + 1). \quad (46)$$

Now, we focus on the remaining expectation in (45). We know that  $p_{cl}^{u1} |\xi_{0cl}|^2 \sim \text{Exp}(1/p_{cl}^{u1} \beta_{0cl})$ , where  $\text{Exp}(x)$  is the exponential distribution with parameter  $x$ . Let  $v_0 \triangleq \sum_{l=1}^L \sum_{c=1}^N p_{cl}^{u1} |\xi_{0cl}|^2$ , which can be written as a single summation, i.e.,  $v_0 = \sum_{k=1}^{NL} \psi_{0k}$ , where  $\psi_{0k} \triangleq p_{cl}^{u1} |\xi_{0cl}|^2$  with  $k = N(l-1) + c$ . Hence,  $\psi_{0k} \sim \text{Exp}(1/\zeta_{0k})$ , where  $\zeta_{0k} \triangleq p_{cl}^{u1} \beta_{0cl}$  with  $k = N(l-1) + c$ . The probability density function of  $v_0$  is given by [30, Theorem 2]

$$f(v_0) = \sum_{h=1}^{\varrho(\mathcal{A}_0)} \sum_{j=1}^{\tau_h(\mathcal{A}_0)} \lambda_{h,j}(\mathcal{A}_0) \frac{\zeta_{0(h)}^{-j}}{(j-1)!} v_0^{j-1} e^{-v_0/\zeta_{0(h)}}, \quad (47)$$

which leads to

$$\mathbb{E} \left\{ \frac{1}{v_0 + 1} \right\} = \sum_{h=1}^{\varrho(\mathcal{A}_0)} \sum_{j=1}^{\tau_h(\mathcal{A}_0)} \lambda_{h,j}(\mathcal{A}_0) \frac{\zeta_{0(h)}^{-j}}{(j-1)!} \int_0^\infty \frac{v_0^{j-1}}{v_0 + 1} e^{-v_0/\zeta_{0(h)}} dv_0, \quad (48)$$

By replacing  $v_0$  with  $-s$ ,

$$\int_0^\infty \frac{v_0^{j-1}}{v_0 + 1} e^{-v_0/\zeta_{0(h)}} dv_0 = (-1)^{j-1} \int_0^\infty \frac{s^{j-1} - 1 + 1}{s - 1} e^{\frac{s}{\zeta_{0(h)}}} ds. \quad (49)$$

With  $s^{j-1} - 1 = (s-1)(s^{j-2} + s^{j-3} + \dots + 1)$ , (49) can be written as

$$\begin{aligned} \int_0^\infty \frac{v_0^{j-1}}{v_0 + 1} e^{-v_0/\zeta_{0(h)}} dv_0 &= (-1)^{j-1} \\ &\times \int_0^\infty (s^{j-2} + s^{j-3} + \dots + 1) e^{\frac{s}{\zeta_{0(h)}}} ds \\ &+ (-1)^{j-1} \int_0^\infty \frac{1}{s-1} e^{\frac{s}{\zeta_{0(h)}}} ds. \end{aligned} \quad (50)$$

Using integration by parts, we can get

$$\int_0^\infty s^m e^{s/a} ds = -(-1)^m a^{m+1} m!, \quad (51)$$

and we also have

$$\int_0^\infty \frac{1}{s-1} e^{s/a} ds = e^{1/a} \mathbf{E}_1(1/a). \quad (52)$$

Substituting (51) and (52) into (50) gives

$$\begin{aligned} \int_0^\infty \frac{v_0^{j-1}}{v_0 + 1} e^{-v_0/\zeta_{0(h)}} dv_0 &= (-1)^{j-1} \\ &\times \left[ - \sum_{m=0}^{j-2} (-1)^m \zeta_{0(h)}^{m+1} m! + e^{1/\zeta_{0(h)}} \mathbf{E}_1(1/\zeta_{0(h)}) \right]. \end{aligned} \quad (53)$$

The theorem follows by substituting (53), (48) and (46) into (45).

## APPENDIX III PROOF OF LEMMA 1

From Jensen's inequality,

$$\log_2 \left( 1 + \frac{1}{\mathbb{E} \left\{ \frac{X}{Y} \right\}} \right) \leq \mathbb{E} \left\{ \log_2 \left( 1 + \frac{X}{Y} \right) \right\} \leq \log_2 \left( 1 + \mathbb{E} \left\{ \frac{X}{Y} \right\} \right). \quad (54)$$

Since  $X$  and  $Y$  are independent,

$$\mathbb{E} \left\{ \frac{X}{Y} \right\} = \mathbb{E} \{X\} \mathbb{E} \left\{ \frac{1}{Y} \right\} \geq \frac{\mathbb{E} \{X\}}{\mathbb{E} \{Y\}}. \quad (55)$$

Utilizing this on the left and right hand side of (54), we can obtain (24).

Comparing (54) and (24), it is obvious that  $\log_2 \left( 1 + \frac{\mathbb{E} \{X\}}{\mathbb{E} \{Y\}} \right)$  lies between the lower and upper bound of  $\mathbb{E} \left\{ \log_2 \left( 1 + \frac{X}{Y} \right) \right\}$ . Therefore, we get the approximation expression in (23).

## APPENDIX IV PROOF OF COROLLARY 1

According to the law of large numbers,

$$\frac{1}{\theta_1} \sum_{i=1}^{\theta_1} x_i - \frac{1}{\theta_1} \sum_{i=1}^{\theta_1} \mathbb{E} \{x_i\} \xrightarrow{a.s.} 0, \text{ as } \theta_1 \rightarrow \infty. \quad (56)$$

Then, since  $\frac{1}{\theta_1} \sum_{i=1}^{\theta_1} \mathbb{E} \{x_i\}$  is bounded away from 0,

$$\frac{1}{\frac{1}{\theta_1} \sum_{i=1}^{\theta_1} x_i} - \frac{1}{\frac{1}{\theta_1} \sum_{i=1}^{\theta_1} \mathbb{E} \{x_i\}} \xrightarrow{a.s.} 0, \text{ as } \theta_1 \rightarrow \infty. \quad (57)$$

That is,

$$\mathbb{E} \left\{ \frac{1}{\frac{1}{\theta_1} \sum_{i=1}^{\theta_1} x_i} \right\} - \frac{1}{\frac{1}{\theta_1} \sum_{i=1}^{\theta_1} \mathbb{E} \{x_i\}} \xrightarrow{a.s.} 0, \text{ as } \theta_1 \rightarrow \infty. \quad (58)$$

Combining (57) and (58), we get

$$\frac{1}{\frac{1}{\theta_1} \sum_{i=1}^{\theta_1} x_i} - \mathbb{E} \left\{ \frac{1}{\frac{1}{\theta_1} \sum_{i=1}^{\theta_1} x_i} \right\} \xrightarrow{a.s.} 0, \text{ as } \theta_1 \rightarrow \infty, \quad (59)$$

which means

$$\theta_1 X - \theta_1 \mathbb{E} \{X\} \xrightarrow{a.s.} 0, \text{ as } \theta_1 \rightarrow \infty. \quad (60)$$

We also have

$$\frac{1}{\theta_2} Y - \frac{1}{\theta_2} \mathbb{E} \{Y\} \xrightarrow{a.s.} 0, \text{ as } \theta_2 \rightarrow \infty. \quad (61)$$

Therefore, since  $\frac{1}{\theta_2} \mathbb{E} \{Y\}$  is also bounded away from 0,

$$\frac{\theta_1 X}{\frac{1}{\theta_2} Y} - \frac{\theta_1 \mathbb{E} \{X\}}{\frac{1}{\theta_2} \mathbb{E} \{Y\}} \xrightarrow{a.s.} 0, \text{ as } \theta_1 \rightarrow \infty \text{ with } \frac{\theta_2}{\theta_1} \rightarrow \omega, \quad (62)$$

which can be written as

$$\omega \theta_1^2 \left( \frac{X}{Y} - \frac{\mathbb{E} \{X\}}{\mathbb{E} \{Y\}} \right) \xrightarrow{a.s.} 0, \text{ as } \theta_1 \rightarrow \infty. \quad (63)$$

By recalling the definition of almost sure convergence, we know that  $\forall \varepsilon > 0$ , there exists  $\theta_0$ , such that for  $\theta_1 > \theta_0$ ,

$$\Pr \left( \left| \omega \theta_1^2 \left( \frac{X}{Y} - \frac{\mathbb{E} \{X\}}{\mathbb{E} \{Y\}} \right) \right| < \varepsilon \right) = 1, \quad (64)$$

where  $\Pr(\cdot)$  denotes probability. Hence,

$$\Pr \left( \left| \frac{X}{Y} - \frac{\mathbb{E} \{X\}}{\mathbb{E} \{Y\}} \right| < \varepsilon \right) = 1, \quad (65)$$

and therefore

$$\frac{X}{Y} - \frac{\mathbb{E} \{X\}}{\mathbb{E} \{Y\}} \xrightarrow{a.s.} 0, \text{ as } \theta_1 \rightarrow \infty. \quad (66)$$

Then, Corollary 1 follows by a direct application of (66).



## APPENDIX V PROOF OF LEMMA 2

Utilizing the property  $\text{tr}(\mathbf{B}_{ll}\mathbf{B}_{ll}^H) = \text{tr}(\mathbf{B}_{ll}^H\mathbf{B}_{ll})$  on (9), we can obtain

$$\mathbb{E}\left\{\text{tr}(\mathbf{B}_{ll}^H\mathbf{B}_{ll})\right\} = 1. \quad (67)$$

Then, substituting (10) into (67) yields  $\alpha_l^2 \mathbb{E}\left\{\text{tr}(\mathbf{G}_{ll}^T\mathbf{G}_{ll}^*)^{-1}\right\} = 1$ , which implies

$$\alpha_l = \sqrt{1/\mathbb{E}\left\{\text{tr}(\mathbf{G}_{ll}^T\mathbf{G}_{ll}^*)^{-1}\right\}}. \quad (68)$$

We know that

$$\mathbb{E}\left\{\text{tr}(\mathbf{G}_{ll}^H\mathbf{G}_{ll})^{-1}\right\} = \text{tr}\left(\mathbb{E}\left\{\left(\mathbf{G}_{ll}^H\mathbf{G}_{ll}\right)^{-1}\right\}\right) \quad (69)$$

$$= \sum_{n=1}^N \mathbb{E}\left\{\left[\left(\mathbf{G}_{ll}^H\mathbf{G}_{ll}\right)^{-1}\right]_{nn}\right\}. \quad (70)$$

From (43), we have

$$\mathbb{E}\left\{\left[\left(\mathbf{G}_{ll}^H\mathbf{G}_{ll}\right)^{-1}\right]_{nn}\right\} = \frac{1}{\beta_{lnl}(M-N)}. \quad (71)$$

Applying this in (69) yields

$$\mathbb{E}\left\{\text{tr}(\mathbf{G}_{ll}^H\mathbf{G}_{ll})^{-1}\right\} = \frac{1}{M-N} \sum_{n=1}^N \frac{1}{\beta_{lnl}}. \quad (72)$$

Therefore, since  $\text{tr}(\mathbf{G}_{ll}^T\mathbf{G}_{ll}^*)^{-1} = \text{tr}(\mathbf{G}_{ll}^H\mathbf{G}_{ll})^{-1}$ ,  $\alpha_l$  is got by applying (72) in (68).

## APPENDIX VI PROOF OF THEOREM 7

Applying Jensen's inequality in (13) yields

$$R_{n0}^{d1} \geq R_{n0}^{d1,L} = \log_2 \left( 1 + \frac{\alpha_0^2 p_{n0}^{d1}}{\sum_{l=1}^L \sum_{c=1}^N p_{cl}^{d1} \mathbb{E}\left\{|\mathbf{g}_{ln0}^T \mathbf{b}_{lcl}|^2\right\} + 1} \right), \quad (73)$$

which can be further written as

$$R_{n0}^{d1,L} = \log_2 \left( 1 + \frac{\alpha_0^2 p_{n0}^{d1}}{\sum_{l=1}^L \sum_{c=1}^N p_{cl}^{d1} \mathbb{E}\left\{\frac{|\mathbf{g}_{ln0}^T \mathbf{b}_{lcl}|^2}{\|\mathbf{b}_{lcl}\|^2} \|\mathbf{b}_{lcl}\|^2\right\} + 1} \right). \quad (74)$$

From the proof of *Theorem 1*, we know that  $\frac{\mathbf{g}_{ln0}^T \mathbf{b}_{lcl}}{\|\mathbf{b}_{lcl}\|}$  has the same distribution as  $\mathbf{g}_{0n0}$  and is independent of  $\mathbf{b}_{lcl}$ . Therefore

$$R_{n0}^{d1,L} = \log_2 \left( 1 + \frac{\alpha_0^2 p_{n0}^{d1}}{\sum_{l=1}^L \sum_{c=1}^N p_{cl}^{d1} \mathbb{E}\left\{\frac{|\mathbf{g}_{ln0}^T \mathbf{b}_{lcl}|^2}{\|\mathbf{b}_{lcl}\|^2}\right\} \mathbb{E}\left\{\|\mathbf{b}_{lcl}\|^2\right\} + 1} \right), \quad (75)$$

and

$$\mathbb{E}\left\{\frac{|\mathbf{g}_{ln0}^T \mathbf{b}_{lcl}|^2}{\|\mathbf{b}_{lcl}\|^2}\right\} = \beta_{ln0}. \quad (76)$$

With  $\|\mathbf{b}_{lcl}\|^2 = \alpha_l^2 \left[\left(\mathbf{G}_{ll}^T\mathbf{G}_{ll}^*\right)^{-1}\right]_{cc}$  and (71), we have

$$\mathbb{E}\left\{\|\mathbf{b}_{lcl}\|^2\right\} = \frac{\alpha_l^2}{\beta_{lcl}(M-N)}. \quad (77)$$

The theorem then follows by substituting (34), (76) and (77) into (75).

## REFERENCES

- [1] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-time Wireless Communications*. Cambridge University Press, 2003.
- [2] D. N. C. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
- [3] H. Jafarkhani, *Space-Time Coding: Theory and Practice*. Cambridge University Press, 2005.
- [4] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Commun.*, pp. 585–595, Nov.-Dec. 1999.
- [5] C.-J. Chen and L.-C. Wang, "Performance analysis of scheduling in multiuser MIMO systems with zero-forcing receivers," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1435–1445, Sept. 2007.
- [6] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, July 2003.
- [7] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 1912–1921, Aug. 2003.
- [8] D. Gesbert, M. Kountouris, R. W. Heath Jr., C.-B. Chae, and T. Salzer, "Shifting the MIMO paradigm," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 36–46, Sept. 2007.
- [9] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2845–2866, June 2010.
- [10] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 528–541, Mar. 2006.
- [11] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.
- [12] T. L. Marzetta, "How much training is required for multiuser MIMO?" in *Proc. Asilomar Conf. Signals, Systems, Comput.*, Oct. 2006, pp. 359–363.
- [13] —, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
- [14] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.
- [15] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.
- [16] A. Pitarokoulis, S. K. Mohammed, and E. G. Larsson, "On the optimality of single-carrier transmission in large-scale antenna systems," *IEEE Wireless Commun. Lett.*, vol. 1, no. 4, pp. 276–279, Aug. 2012.
- [17] S. Wagner, R. Couillet, D. T. M. Slock, and M. Debbah, "Large system analysis of zero-forcing precoding in MISO broadcast channels with limited feedback," in *Proc. IEEE Int. Work. Signal Process. Adv. Wireless Commun. (SPAWC)*, June 2010.
- [18] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "The multicell multiuser MIMO uplink with very large antenna arrays and a finite-dimensional channel," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2350–2361, June 2013.
- [19] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Jan. 2013.
- [20] F. Fernandes, A. Ashikhmin, and T. L. Marzetta, "Inter-cell interference in noncooperative TDD large scale antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 192–201, Feb. 2013.
- [21] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.
- [22] K. Appaiah, A. Ashikhmin, and T. L. Marzetta, "Pilot contamination reduction in multi-user TDD systems," in *Proc. IEEE International Conference on Communication (ICC)*, May 2010, pp. 1–5.
- [23] Q. Zhang, S. Jin, K.-K. Wong, H. Zhu, and M. Matthaiou, "Power scaling of uplink massive MIMO systems with arbitrary-rank channel means," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 966–981, Oct. 2014.
- [24] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless link?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.
- [25] H. Yang and T. L. Marzetta, "Performance of conjugate and zero-forcing beamforming in large-scale antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 172–179, Feb. 2013.

- [26] H. Q. Ngo, T. Q. Duong, and E. G. Larsson, "Uplink performance analysis of multicell MU-MIMO with zero-forcing receivers and perfect CSI," in *Proc. Commun. Technologies Workshop (SWE-CTW)*, Oct. 2011, pp. 40–45.
- [27] X. Li, S. Jin, X. Gao, and M. R. McKay, "Capacity bounds and low complexity transceiver design for double-scattering MIMO multiple access channels," *IEEE Trans. Signal Process.*, vol. 58, no. 5, pp. 2809–2822, May 2010.
- [28] H. Shin and M. Z. Win, "MIMO diversity in the presence of double scattering," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 2976–2996, July 2008.
- [29] D. Gore, R. W. Heath Jr., and A. Paulraj, "Transmit selection in spatial multiplexing systems," *IEEE Commun. Lett.*, vol. 6, no. 11, pp. 491–493, Nov. 2002.
- [30] A. Bletsas, H. Shin, and M. Z. Win, "Cooperative communications with outage-optimal opportunistic relaying," *IEEE Trans. Wireless Commun.*, vol. 6, no. 9, pp. 3450–3460, Sept. 2007.